

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/572,711
Source: IFwo
Date Processed by STIC: 3/31/06

ENTERED



IFWP

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/572,711

DATE: 03/31/2006

TIME: 12:11:03

Input Set : A:\11916.0059.PCUS01.ST25.txt
 Output Set: N:\CRF4\03292006\J572711.raw

3 <110> APPLICANT: Bogosian, Gregg
 4 O'Neill, Julia P.
 5 Smith, Hong Q.
 7 <120> TITLE OF INVENTION: Prevention of Incorporation of Non-Standard Amino
 Acids into
 8 Protein
 10 <130> FILE REFERENCE: 11916.0059.PCUS01
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/572,711
 C--> 12 <141> CURRENT FILING DATE: 2006-03-20
 12 <150> PRIOR APPLICATION NUMBER: PCT/US 2004/031224
 13 <151> PRIOR FILING DATE: 2004-09-23
 15 <150> PRIOR APPLICATION NUMBER: US 60/505,807
 16 <151> PRIOR FILING DATE: 2003-09-25
 18 <160> NUMBER OF SEQ ID NOS: 16
 20 <170> SOFTWARE: PatentIn version 3.3
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 1344
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Escherichia coli
 27 <400> SEQUENCE: 1
 28 atggatcaga catattctct ggagtcatc ctaaccatg tccaaaagcg cgaccgaaat 60
 30 caaaccgagt tcgcgcaagc cgttcgtgaa gtaatgacca cactctggcc ttttcttgaa 120
 32 caaaaatccaa aatatcgcca gatgtcatta ctggagcgctc tgggtgaacc ggagcgctg 180
 34 atccagttc gcgtggatag ggttgcgtatcg cgcaaccaga tacaggtcaa ccgtgcattgg 240
 36 cgtgtcagtc tcaagctctgc catcgcccg tacaaggcg gtatgcgtt ccatccgtca 300
 38 gttaacctt ccattctcaa attcctcgcc tttgaacaaa cttcaaaaaa tgccctgact 360
 40 actctgccga tgggcgggtgg taaaggcgcc agcgatttcg atccgaaagg aaaaagcgaa 420
 42 ggtgaagtga tgcgttttg ccaggcgctg atgactgaac tttatcgcc a cctggcg 480
 44 gataccgacg ttccggcagg tggatcgcc gttgggtggc gtgaagtcgg ctttatggcg 540
 46 gggatgatga aaaagctctc caacaatacc gcctgcgtt tcaccggtaa gggccttca 600
 48 tttggcggca gtcttattcg cccggaaatcg accggctacg gtctggtttta tttcacagaa 660
 50 gcaatgctaa aacgccacgg tatgggtttt gaaggatgc gcgttccgt ttctggctcc 720
 52 ggcaacgtcg cccagtcgc tatcgaaaaa gcgatggaaat ttgggtctcg tttatggcg 780
 54 gcgtcagact ccagcggcac tggatgttgc gaaaggcgat tcacgaaaga gaaactggca 840
 56 cgtttatcg aaatcaaagc cagccgcgt ggtcgatgg cagattacgc caaagaattt 900
 58 ggtctggct atctcgaaagg ccaacagccg tggctctac cggttgatcg cgcctgcct 960
 60 tgcgccaccc agaatgaact ggtatgttgc gccgcgcac agttatcg tttatggcg 1020
 62 aaagccgtcg ccgaaggggc aaatatcgcc accaccatcg aagcgactga actgttccag 1080
 64 caggcaggcg tactatttcg accgggtaaa gcccgtatcg ctgggtggcgt cgctacatcg 1140
 66 ggcctggaaa tggcacaaaaa cgctgcgcgc ctgggtggaa aagccgagaa agttgacgca 1200
 68 cgtttgcattc acatcatgtt ggtatccac catgcctgtt ttgagcatgg tggtaaggt 1260
 70 gagcaaacca actacgtgca gggcgcgaac attgcccgtt ttgtgaaggt tgccgatgcg 1320
 72 atgctggcgc aggggtgtat ttaa 1344
 75 <210> SEQ ID NO: 2

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/572,711

DATE: 03/31/2006
 TIME: 12:11:03

Input Set : A:\11916.0059.PCUS01.ST25.txt
 Output Set: N:\CRF4\03292006\J572711.raw

76 <211> LENGTH: 447
 77 <212> TYPE: PRT
 78 <213> ORGANISM: Escherichia coli
 80 <400> SEQUENCE: 2
 82 Met Asp Gln Thr Tyr Ser Leu Glu Ser Phe Leu Asn His Val Gln Lys
 83 1 5 10 15
 86 Arg Asp Pro Asn Gln Thr Glu Phe Ala Gln Ala Val Arg Glu Val Met
 87 20 25 30
 90 Thr Thr Leu Trp Pro Phe Leu Glu Gln Asn Pro Lys Tyr Arg Gln Met
 91 35 40 45
 94 Ser Leu Leu Glu Arg Leu Val Glu Pro Glu Arg Val Ile Gln Phe Arg
 95 50 55 60
 98 Val Val Trp Val Asp Asp Arg Asn Gln Ile Gln Val Asn Arg Ala Trp
 99 65 70 75 80
 102 Arg Val Gln Phe Ser Ser Ala Ile Gly Pro Tyr Lys Gly Gly Met Arg
 103 85 90 95
 106 Phe His Pro Ser Val Asn Leu Ser Ile Leu Lys Phe Leu Gly Phe Glu
 107 100 105 110
 110 Gln Thr Phe Lys Asn Ala Leu Thr Thr Leu Pro Met Gly Gly Gly Lys
 111 115 120 125
 114 Gly Gly Ser Asp Phe Asp Pro Lys Gly Lys Ser Glu Gly Glu Val Met
 115 130 135 140
 118 Arg Phe Cys Gln Ala Leu Met Thr Glu Leu Tyr Arg His Leu Gly Ala
 119 145 150 155 160
 122 Asp Thr Asp Val Pro Ala Gly Asp Ile Gly Val Gly Arg Glu Val
 123 165 170 175
 126 Gly Phe Met Ala Gly Met Met Lys Lys Leu Ser Asn Asn Thr Ala Cys
 127 180 185 190
 130 Val Phe Thr Gly Lys Gly Leu Ser Phe Gly Gly Ser Leu Ile Arg Pro
 131 195 200 205
 134 Glu Ala Thr Gly Tyr Gly Leu Val Tyr Phe Thr Glu Ala Met Leu Lys
 135 210 215 220
 138 Arg His Gly Met Gly Phe Glu Gly Met Arg Val Ser Val Ser Gly Ser
 139 225 230 235 240
 142 Gly Asn Val Ala Gln Tyr Ala Ile Glu Lys Ala Met Glu Phe Gly Ala
 143 245 250 255
 146 Arg Val Ile Thr Ala Ser Asp Ser Ser Gly Thr Val Val Asp Glu Ser
 147 260 265 270
 150 Gly Phe Thr Lys Glu Lys Leu Ala Arg Leu Ile Glu Ile Lys Ala Ser
 151 275 280 285
 154 Arg Asp Gly Arg Val Ala Asp Tyr Ala Lys Glu Phe Gly Leu Val Tyr
 155 290 295 300
 158 Leu Glu Gly Gln Gln Pro Trp Ser Leu Pro Val Asp Ile Ala Leu Pro
 159 305 310 315 320
 162 Cys Ala Thr Gln Asn Glu Leu Asp Val Asp Ala Ala His Gln Leu Ile
 163 325 330 335
 166 Ala Asn Gly Val Lys Ala Val Ala Glu Gly Ala Asn Met Pro Thr Thr
 167 340 345 350
 170 Ile Glu Ala Thr Glu Leu Phe Gln Ala Gly Val Leu Phe Ala Pro

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/572,711

DATE: 03/31/2006
TIME: 12:11:03

Input Set : A:\11916.0059.PCUS01.ST25.txt
Output Set: N:\CRF4\03292006\J572711.raw

171 355 360 365
 174 Gly Lys Ala Ala Asn Ala Gly Gly Val Ala Thr Ser Gly Leu Glu Met
 175 370 375 380
 178 Ala Gln Asn Ala Ala Arg Leu Gly Trp Lys Ala Glu Lys Val Asp Ala
 179 385 390 395 400
 182 Arg Leu His His Ile Met Leu Asp Ile His His Ala Cys Val Glu His
 183 405 410 415
 186 Gly Gly Glu Gly Glu Gln Thr Asn Tyr Val Gln Gly Ala Asn Ile Ala
 187 420 425 430
 190 Gly Phe Val Lys Val Ala Asp Ala Met Leu Ala Gln Gly Val Ile
 191 435 440 445
 194 <210> SEQ ID NO: 3
 195 <211> LENGTH: 1344
 196 <212> TYPE: DNA
 197 <213> ORGANISM: Escherichia coli
 199 <400> SEQUENCE: 3
 200 atgatcaga catattctct ggagtcatc ctcaaccatg tccaaaagcg cgaccgaaat 60
 202 caaaccgagt tcgcgcaga cgttcgtgaa gtaatgacca cactctggcc ttttcttgaa 120
 204 caaaatccaa aatatcgcca gatgtcatta ctggagcgctc tgggtgaacc ggagcgcgtg 180
 206 atccagtttc gcgtggatg ggttgcgtat cgcaaccaga tacaggtcaa ccgtgcattgg 240
 208 cgtgtgcagt tcagctctgc catcgcccg tacctggcg gtatgcgcctt ccattccgtca 300
 210 gttAACCTTT ccattctcaa attcctcgcc tttgaacaaa ctttcaaaaaa tgccctgact 360
 212 actctgcgcga tgggcgttgg taaaggcgcc agcgatttcg atccgaaagg aaaaagcgaa 420
 214 ggtgaagtga tgcgttttg ccaggcgctg atgactgaac tttatcgcca cctggcgccg 480
 216 gataccgacg ttccggcagg tggatcgccg gttgggtggc gtgaagtcgg ctttatggcg 540
 218 gggatgatga aaaagcttc caacaatacc gcctgcgtct tcaccggtaa gggccttca 600
 220 tttggcggca gtcttattcg cccggaaagct accggctacg gtctggttt tttcacagaa 660
 222 gcaatgctaa aacgcccacgg tatgggtttt gaagggtatgc gcggttccgt ttctggctcc 720
 224 ggcaacgtcg cccagttacgc tatcgaaaaa gcatggat ttgggtgcgc tggatcact 780
 226 gcgtcagact ccagcggcac tggatgtat gaaaggcgat tcacgaaaga gaaactggca 840
 228 cgctttatcg aaatcaaaggc cagccgcgt ggtcgagtgg cagattacgc caaagaattt 900
 230 ggtctggctc atctcgaaagg ccaacagccg tggatcgatc cggttgcgtat cgccctgcct 960
 232 tgcgccaccc agaatgaaact ggtatgtac gcccgcgc acgttatcg taatggcg 1020
 234 aaagccgtcg ccgaaggggc aaatatgccc accaccatcg aagcgactga actgttccag 1080
 236 caggcaggcg tactattgc accgggtaaa gcccgtatcg ctgggtggcg cgctacatcg 1140
 238 ggcctggaaa tggcacaaaa cgctgcgcgc ctgggtggaa aagccgagaaa agttgacgca 1200
 240 cgtttgcatac acatcatgct ggtatccac catgcctgtt ttgagcatgg tggtaaggt 1260
 242 gagcaaacca actacgtgca gggcgcgaac attgcccgtt ttgtgaaggt tgccgatgcg 1320
 244 atgctggcgca agggtgtat ttaa 1344
 247 <210> SEQ ID NO: 4
 248 <211> LENGTH: 447
 249 <212> TYPE: PRT
 250 <213> ORGANISM: Escherichia coli
 252 <400> SEQUENCE: 4
 254 Met Asp Gln Thr Tyr Ser Leu Glu Ser Phe Leu Asn His Val Gln Lys
 255 1 5 10 15
 258 Arg Asp Pro Asn Gln Thr Glu Phe Ala Gln Ala Val Arg Glu Val Met
 259 20 25 30
 262 Thr Thr Leu Trp Pro Phe Leu Glu Gln Asn Pro Lys Tyr Arg Gln Met

RAW SEQUENCE LISTING DATE: 03/31/2006
 PATENT APPLICATION: US/10/572,711 TIME: 12:11:03

Input Set : A:\11916.0059.PCUS01.ST25.txt
 Output Set: N:\CRF4\03292006\J572711.raw

263	35	40	45
266	Ser Leu Leu Glu Arg Leu Val Glu Pro Glu Arg Val Ile Gln Phe Arg		
267	50	55	60
270	Val Val Trp Val Asp Asp Arg Asn Gln Ile Gln Val Asn Arg Ala Trp		
271	65	70	75
274	80	Arg Val Gln Phe Ser Ser Ala Ile Gly Pro Tyr Leu Gly Gly Met Arg	
275	85	90	95
278	Phe His Pro Ser Val Asn Leu Ser Ile Leu Lys Phe Leu Gly Phe Glu		
279	100	105	110
282	Gln Thr Phe Lys Asn Ala Leu Thr Thr Leu Pro Met Gly Gly Gly Lys		
283	115	120	125
286	Gly Gly Ser Asp Phe Asp Pro Lys Gly Lys Ser Glu Gly Glu Val Met		
287	130	135	140
290	Arg Phe Cys Gln Ala Leu Met Thr Glu Leu Tyr Arg His Leu Gly Ala		
291	145	150	155
294	160	165	170
295	175	170	175
298	Gly Phe Met Ala Gly Met Met Lys Lys Leu Ser Asn Asn Thr Ala Cys		
299	180	185	190
302	305	310	315
303	Val Phe Thr Gly Lys Gly Leu Ser Phe Gly Gly Ser Leu Ile Arg Pro		
306	320	325	330
307	Glu Ala Thr Gly Tyr Gly Leu Val Tyr Phe Thr Glu Ala Met Leu Lys		
310	335	340	345
311	350	355	360
314	365	370	375
315	380	385	390
318	395	400	405
319	410	415	420
322	425	430	
323			
326			
327			
330			
331			
334			
335			
338			
339			
342			
343			
346			
347			
350			
351			
354			
355			
358			
359			

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/572,711

DATE: 03/31/2006
 TIME: 12:11:03

Input Set : A:\11916.0059.PCUS01.ST25.txt
 Output Set: N:\CRF4\03292006\J572711.raw

362 Gly Phe Val Lys Val Ala Asp Ala Met Leu Ala Gln Gly Val Ile
 363 435 440 445
 366 <210> SEQ ID NO: 5
 367 <211> LENGTH: 1101
 368 <212> TYPE: DNA
 369 <213> ORGANISM: *Bacillus cereus*
 371 <400> SEQUENCE: 5
 372 atgacattag aaatcttcga atacttagaa aaatatgatt atgagcaagt agtattttgt 60
 374 caagataaaag aatctggttt aaaagcaatt attgcaattc atgataacaac acttggaccg 120
 376 gctcttggtg gaacaagaat gtggacatat gattctgaag aagcggcgat tgaagatgca 180
 378 ttgcgtcttg caaaaggat gacataaaaa aacgcagcag ctggttaaa cttaggtgg 240
 380 gcgaaaacag taattatcgg tgatcctcgtaa aagataaga gcgaaagcaat gttccgtgca 300
 382 ctaggacgtt atatccaagg actaaacgga cgttacatttta cagctgaaga tggtggtaca 360
 384 acagtagatg atatggatat tatccatgaa gaaactgact ttgttaacagg tatctcacca 420
 386 tcattcgggtt cttctggtaa cccatctccg gtaactgcat acgggtttt ccgtggatgt 480
 388 aaagcagctg caaaagaagc tttcgtact gacaatttag aaggaaaagt aattgtctgtt 540
 390 caaggcgttg gtaacgttagc atatcaccta tgcaaacatt tacacgctga aggagcaaaa 600
 392 ttaatcgtta cagatattaa taaagaagct gtacaacgtg ctgttagaaga attcggtgca 660
 394 tcagcagttg aaccaaataa aatttacggt gttgaatgcg atatttacgc accatgtgca 720
 396 ctaggcgcaaa cagttaatga tgaaactatt ccacaactta aagcaaaagt aatcgcaggt 780
 398 tctgcaata accaattaaa agaagatcgt catggtgaca tcattcatga aatgggtatt 840
 400 gtatacgcac cagattatgt aattaatgca ggtggcgtaa ttaacgttagc agacgaatta 900
 402 tatggatatac atagagaacg tgactaaaaa cgtgtttagt ctatttatga cacgattgca 960
 404 aaagtaatcg aaatttcaaa acgcgtggc atagcaactt atgttagcggc agatcgtcta 1020
 406 gctgaagagc gcattgcaag cttgaagaat tctcgttagca cttacttacg caacggtcac 1080
 408 gatattatta gccgtcgcta a 1101
 411 <210> SEQ ID NO: 6
 412 <211> LENGTH: 366
 413 <212> TYPE: PRT
 414 <213> ORGANISM: *Bacillus cereus*
 416 <400> SEQUENCE: 6
 418 Met Thr Leu Glu Ile Phe Glu Tyr Leu Glu Lys Tyr Asp Tyr Glu Gln
 419 1 5 10 15
 422 Val Val Phe Cys Gln Asp Lys Glu Ser Gly Leu Lys Ala Ile Ile Ala
 423 20 25 30
 426 Ile His Asp Thr Thr Leu Gly Pro Ala Leu Gly Gly Thr Arg Met Trp
 427 35 40 45
 430 Thr Tyr Asp Ser Glu Glu Ala Ala Ile Glu Asp Ala Leu Arg Leu Ala
 431 50 55 60
 434 Lys Gly Met Thr Tyr Lys Asn Ala Ala Gly Leu Asn Leu Gly Gly
 435 65 70 75 80
 438 Ala Lys Thr Val Ile Ile Gly Asp Pro Arg Lys Asp Lys Ser Glu Ala
 439 85 90 95
 442 Met Phe Arg Ala Leu Gly Arg Tyr Ile Gln Gly Leu Asn Gly Arg Tyr
 443 100 105 110
 446 Ile Thr Ala Glu Asp Val Gly Thr Thr Val Asp Asp Met Asp Ile Ile
 447 115 120 125
 450 His Glu Glu Thr Asp Phe Val Thr Gly Ile Ser Pro Ser Phe Gly Ser
 451 130 135 140

VERIFICATION SUMMARY DATE: 03/31/2006
PATENT APPLICATION: US/10/572,711 TIME: 12:11:04

Input Set : A:\11916.0059.PCUS01.ST25.txt
Output Set: N:\CRF4\03292006\J572711.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date